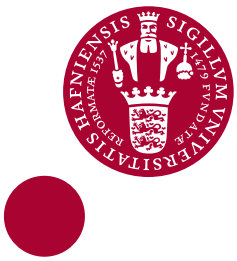

Referral population studies underestimate differences between human papillomavirus assays in primary cervical screening

Rebolj, M.; Njor, S.; Lynge, E.; Preisler, S.; Ejegod, D.; Rygaard, C.; Bonde, J.

Published in:
Cytopathology

DOI:
10.1111/cyt.12451

Publication date:
2017

Document version
Publisher's PDF, also known as Version of record

Accepted: 3 May 2017

DOI: 10.1111/cyt.12451

ORIGINAL ARTICLE

WILEY

Referral population studies underestimate differences between human papillomavirus assays in primary cervical screening

M. Rebolj[1–3] | S. Njor[1,4] | E. Lynge[1] | S. Preisler[2,3] | D. Ejegod[3] | C. Rygaard[3] | J. Bonde[2,3]

[1]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[2]Clinical Research Centre, Copenhagen University Hospital, Hvidovre, Denmark

[3]Department of Pathology, Copenhagen University Hospital, Hvidovre, Denmark

[4]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

Correspondence
Matejka Rebolj, Wolfson Institute of Preventive Medicine, Centre for Cancer Prevention, Queen Mary University of London, London, UK.
Email: m.rebolj@qmul.ac.uk

Funding information
Genomica, the manufacturer of the CLART HPV2 assay; Hologic, the manufacturer of the APTIMA HPV assay; Roche, the manufacturer of the cobas HPV test; and Qiagen, the manufacturer of the Hybrid Capture 2 assay, provided tests, instrumentation, and limited co-funding for laboratory material. M.R. and S.P. were funded by the Danish Strategic Research Council (grant number: 10-092793). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The researchers worked independently of the funders.

Objective: We studied how representative cytologically abnormal women ("referral populations") are with respect to uncovering differences between human papillomavirus (HPV) assays in the primary screening where most women are cytologically normal.

Methods: A total of 4997 women were tested with SurePath® cytology, and Hybrid Capture 2 (HC2), cobas, CLART and APTIMA HPV assays. Women with positive test results were offered a follow-up. For all detected HPV infections and HPV-positive high-grade cervical intraepithelial neoplasia (≥CIN2), we studied the distributions of assay-specific signal strengths in the baseline samples as documented by the assays' automatically generated reports. We calculated the likelihood of test result discordance as the proportion of HPV-positive samples that were not confirmed by all four assays.

Results: Median signal strengths for HPV infections were weaker in normal than abnormal cytology ($P<.001$, adjusted for women's age, multiple infections and the reason for taking the sample). For HC2, they were RLU/CO 11.0 (interquartile range, IQR: 3.3-52.8) vs 124.2 (IQR: 22.8-506.9), respectively; for cobas, Ct 33.5 (IQR: 29.6-37.5) vs 26.9 (IQR: 23.7-31.3), respectively; for APTIMA, S/CO 10.2 (IQR: 5.8-11.3) vs 11.1 (IQR: 9.4-15.5), respectively. Similar patterns were observed for HPV-positive ≥CIN2. The four HPV assays more frequently returned discordant test results in normal than in abnormal cytology. Relative frequency of discordance in detecting HPV infections was 0.39 (95% confidence interval: 0.33-0.48) for abnormal vs normal cytology.

Conclusions: These data suggest that referral population studies, by not including sufficient numbers of cytology normal women, underestimate the differences between HPV assays that would become apparent in primary screening.

KEYWORDS
assays, cervical cancer, cytology, human papillomavirus, mass screening

1 | INTRODUCTION

Human papillomavirus (HPV) testing is going to replace cytology in primary cervical screening. Randomised trials comparing the two testing modalities in screening showed that women with negative HPV test results have a lower risk of developing cervical cancer than women with normal cytology.[1]

Cervical screening laboratories can now choose between more than 100 commercially available HPV assays. The randomised trials were, however, undertaken either with one commercially available HPV assay, Hybrid Capture 2 (HC2) or with an in-house polymerase chain reaction (PCR) assay using GP5+/6+ primers. Other, newer, HPV assays will most likely not be submitted to similarly rigorous but costly and time-consuming randomised trials. Instead, their accuracy of detecting high-grade cervical intraepithelial neoplasia (CIN), a treatable screening endpoint, has been compared to that of HC2 or GP5+/6+ in several smaller, predominantly split-sample, studies.[2–20]

The caveat is that split-sample testing literature is dominated by studies of referral populations.[8,9,11–20] As referral populations typically include women with abnormal cervical cytology who have a high risk of ≥CIN2, such studies have several practical advantages over primary screening studies. The high risk of ≥CIN2 means that adequately powered referral population studies can be smaller than adequately powered primary screening studies where the prevalence of ≥CIN2 is much lower. Furthermore, women with abnormal cytology are routinely offered a follow-up, so CIN lesions detected through abnormal cytology can be traced in routine pathology registrations. In contrast, follow-up of cytology-normal/HPV-positive women at present necessitates ethical approval and an additional research infrastructure.

In primary screening, most HPV-positive women have normal cytology, and the main reason for replacing cytology with HPV testing is to detect also the CIN missed by cytology (ie, the cytologically normal high-grade CIN). Referral populations, therefore, represent only selected subgroups of women undergoing primary screening, and it remains unknown whether they are an adequate substitute for primary screening populations when HPV assays are being compared.

To shed light on the use of referral populations in studies evaluating HPV assays for primary screening, we compared HPV infections and ≥CIN2 detected by four commercially available HPV assays in women with normal vs abnormal cytology.

2 | METHODS

2.1 | Study design

We used data from the Horizon study, the design of which was described in detail previously.[2,21–28] In short, we collected consecutive, routine, SurePath® samples from Department of Pathology of Copenhagen University Hospital Hvidovre in June-August 2011. Samples from 5034 women were tested with liquid-based cytology (LBC) and four HPV assays: HC2 (QIAGEN, Gaithersburg, MD, USA), cobas (Roche, Pleasanton, CA, USA), CLART (Genomica, Madrid, Spain) and APTIMA (Hologic, San Diego, CA, USA). HC2 was tested on post-quot cytology material. The residual original material was diluted 1:1 in SurePath for aliquoting into tubes for cobas, APTIMA, and CLART testing (Appendix). Cytology was read routinely following the Bethesda 2001 criteria, blinded to HPV testing outcomes. Abnormal cytology was defined as atypical squamous cells of undetermined significance (ASCUS) or worse.

International Agency for Research on Cancer (IARC) classifies 13 genotypes as high-risk (oncogenic): 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68. Positive HPV test results on HC2, cobas and APTIMA were defined as intended by the respective manufacturers, and signal strengths for positive baseline HPV test results were determined from the assays' automatically generated testing documentation. By design, HC2 detects the 13 high-risk genotypes in combination, with a signal threshold of ≥1.0 relative light units per cut-off (RLU/CO). Cobas, on the other hand, detects genotypes 16 and 18 in two separate channels (signal thresholds: critical threshold (Ct) values ≤40.5 and ≤40.0) and the remaining 11 high-risk plus one possibly carcinogenic genotype (genotype 66) in combination in a third channel (signal threshold: ≤40.0 Ct). APTIMA detects the 13 high-risk genotypes plus genotype 66 in combination, with a signal threshold of ≥0.5 signal to a cut-off (S/CO) value. In these three assays, low-risk genotypes return a negative test result (except in cases of cross-reactivity, described in detail previously).[28]

Among the four studied assays, CLART was the only assay that allowed individual detection of 35 HPV genotypes (including all 13 defined as high-risk by IARC) and was used in the analysis to discriminate between single and multiple infections. Within the context of primary cervical screening, we considered samples to have a positive CLART test result if the assay detected one or more of the 13 high-risk HPV genotypes. All other infections detected by CLART were considered to represent negative, non-actionable, screening findings. In the analysis, these infections could only contribute to multiple infections in the presence of at least one high-risk genotype.

Women with abnormal cytology were managed according to routine Danish recommendations (Appendix). Women with cytology-normal/HPV-positive test results on one or more HPV assays were invited, for study purposes, for repeated cytology and HPV testing in 18 months. All colposcopies were undertaken under routine conditions either by a hospital or privately practicing gynaecologists. In Denmark, it is recommended to take directed biopsies from all suspicious areas after application of acetic acid, and a random biopsy from all quadrants if lesions are not visible. The most severe follow-up testing results in 2.5 years after the baseline were determined through linkage to the national Danish Pathology Register (Patobank).[29]

2.2 | Statistical analysis

Samples with inadequate cytology (N=25, 0.5%) were excluded, as were samples with invalid CLART test results (N=12, 0.2%). The remaining 4997 samples were included in the analysis. They represent a typical collection of samples handled by a cytology laboratory

during its routine operations. Their screening histories were determined from the Patobank from 1 January 2000 until the baseline testing date. These screening histories were used to determine the reason for taking each sample. The reasons were categorised as either primary screening or follow-up of recent abnormalities (see Appendix for detailed definitions). In the present analysis, we limited the primary screening population to samples taken at age 30-65 years (N=2846 out of 4997), as younger women are not considered for HPV-based primary screening. Approximately 96% of these women had normal, and 4% had abnormal routine cytology. Previous studies used various definitions for their referral populations. Typically, they included women with cytological abnormalities at any age and regardless of their screening history, ranging from women with single ASCUS samples to women attending colposcopy owing to high-grade cytological abnormalities or several consecutive low-grade screening-detected abnormalities. Our primary definition of a referral population included all women with ≥ASCUS, so as not to exaggerate the differences between our referral and screening populations (N=367 out of 4997, of which 127 were also included in the primary screening population described above). We separately investigated women with follow-up samples in the Horizon study showing normal baseline cytology (N=516 out of 4997); samples from these women are also readily available from routine pathology registrations and are sometimes included among referral populations.

We calculated median signal strengths and their interquartile ranges (IQR) for HC2, cobas and APTIMA. The dispersion of the signal strength values was estimated by the range between IQR[1]/median and IQR[3]/median. If more than one cobas' channel returned a positive test result (18% of cobas-positive women with normal and 26% with abnormal cytology), we considered the signal strength on the most strongly positive channel (=the lowest Ct value). Signals closer to the cut-off were considered to be weaker. Signal strength is not available for CLART, so signal strength statistics could not be determined.

Differences in median signal strengths of the detected HPV infections and the detected ≥CIN2 lesions were assessed between different populations using lognormal distribution. Lognormal distribution was used for testing as it provided the best fit for the data. *P*-values based on other distributions were very close to the reported *P*-values based on the lognormal distribution. They were adjusted for all known characteristics of the woman and/or the study sample; see table footnotes for details specific to each comparison. We compared the following populations: (1) abnormal vs normal cytology in all samples (N=4997), (2) abnormal vs normal cytology in primary screening samples (N=2846), (3) primary screening samples (N=2864) vs referral population samples with abnormal cytology (N=367) and (4) primary screening samples with normal cytology (N=2719) vs referral population samples with normal cytology (N=516). Analyses were performed with SAS Ver. 9.3 (SAS Institute, Cary, NC, USA).

Finally, we evaluated the differences in the assays' ability to detect HPV infections and ≥CIN2 by determining the degree of their concordance in baseline testing results. As previously,[2] concordance was calculated as the conditional probability that all four assays returned a positive test result if at least one of the four was positive, and discordance as [100%-concordance]. The 95% confidence intervals (CI) for relative discordance comparing the various populations were calculated assuming lognormal distribution.

2.3 | Ethical approval

Baseline testing on the residual material was undertaken as a quality development study and did, in line with the Danish regulation of biomedical research, not require ethical approval. Ethical Committee of the Danish Capital Region approved the follow-up of cytology-normal/HPV-positive women (H-4-2012-120), and women signed informed consent. The study was notified to the Danish Data Inspection Agency (notification numbers 2010-41-5594 and AHH-2015-080/I-Suite: 04109).

3 | RESULTS

3.1 | Sample description

Among all 4997 samples, 367 (7%) were associated with abnormal, and 4630 (93%) with normal baseline cytology (Table 1). In total, 1660 (33%) samples showed HPV infections detected by one or more assays and 175 (4%) were associated with HPV-positive ≥CIN2.

In primary screening at 30-65 years (N=2846), 127 (4%) samples were associated with abnormal and 2719 (96%) with normal cytology; 651 (23%) samples showed HPV infections and 50 (2%) were associated with HPV-positive ≥CIN2.

3.2 | Signal strengths in detected HPV infections

A comparison of central tendency measures showed that signal strengths, particularly for HC2, were not distributed normally (Table 2).

Median signal strengths for positive HPV test results in all 4997 samples were weaker in normal than in abnormal cytology (Table 3). For HC2, the median RLU/CO in normal cytology was 11.0 (IQR: 3.3-52.8), compared to 124.2 (IQR: 22.8-506.9, $P \leq .001$) with abnormal cytology. For cobas, the median Ct values were 33.5 (IQR: 29.6-37.5) and 26.9 (IQR: 23.7-31.3, $P \leq .001$), respectively, and for APTIMA, the median S/CO values were 10.2 (IQR: 5.8-11.3) and 11.1 (IQR: 9.4-15.5, $P \leq .001$), respectively.

For HC2, the ranges for IQR/median ratios were 0.3[IQR1]-4.8[IQR3] for normal and 0.2[IQR1]-4.1[IQR3] for abnormal cytology; for cobas, they were 0.9[IQR1]-1.1[IQR3] for normal and 0.9[IQR1]-1.2[IQR3] for abnormal cytology; and for APTIMA, they were 0.6[IQR1]-1.1[IQR3] for normal and 0.8[IQR1]-1.4[IQR3] for abnormal cytology (not tabulated). This suggested that, in relative terms, HPV signal strengths were approximately as heterogeneous for normal as for abnormal cytology.

In primary screening, 558 (86%) of 651 samples with positive HPV test results on one or more assays had normal, and 93 (14%) had abnormal baseline cytology (Table 1/footnote). The patterns in

TABLE 1 Age distribution, cytology test results, detected HPV infections and HPV-positive ≥CIN2 lesions in all women and primary screening at age 30-65 y

	All women (N=4997) (%)[b]	Primary screening at age 30-65 (N=2846) (%)
Age (years)		
<23	161 (3)	–
23-29	1514 (30)	–
30-65	3212 (64)	2846 (100)[d]
>65	110 (2)	–
Cytology at baseline		
Normal	4630 (93)	2719 (96)
Abnormal (≥ASCUS)	367 (7)[c]	127 (4)
Endpoints detected by ≥1 HPV assay		
HPV infections	1660 (33)	651 (23)[e]
HPV-positive ≥CIN2[a]	175 (4)	50 (2)

ASCUS, atypical squamous cells of undetermined significance; CIN, cervical intraepithelial neoplasia; HPV, Human Papillomavirus.

[a]Lesions positive on at least one of the four HPV assays. Two additional cases of ≥CIN2 had a negative test result on all four HPV assays (one of which in a woman with a primary screening sample at 30-65 y). As the focus of this analysis was on ≥CIN2 detectable by HPV assays, these two cases were not considered here.

[b]Of the 4997 women with consecutive, routine samples, 2846 (57%) had primary screening samples at age 30-65 y, 1509 (30%) had primary screening samples at other ages, and the remaining 642 (13%) women had samples taken for follow-up of recent abnormalities. Of the 642 women with follow-up samples, 516 (80%) had normal cytology and the remaining 126 (20%) had abnormal cytology. Of the 367 women with abnormal cytology, 127 (35%) were from primary screening at age 30-65 y, 114 (31%) from primary screening at other ages and the remaining 126 (34%) were from follow-up of recent abnormalities.

[c]These women were described in detail by Rebolj et al.[25]

[d]These women were described in detail by Rebolj et al.[27] The present analysis excluded women with inadequate cytology and/or an invalid CLART test result at baseline.

[e]Of the 651 women with a positive HPV test result on at least one assay, 558 (86%) had normal and 93 (14%) had abnormal cytology.

signal strengths that were observed for all 4997 samples when comparing normal vs abnormal cytology (Table 3 above) were also seen in primary screening, regardless of the assay (Table 3 below).

Next, we compared the complete screening population (N=2846; of which 4% of samples were associated with abnormal cytology) with the referral population with abnormal cytology (N=367). The median signal strengths of HPV-positive samples were weaker in the primary screening than in the referral population ($P<.001$; Table 4).

In the referral population with normal cytology (N=516), the median signal strengths (Table 5) appeared similar to those in samples associated with normal cytology from primary screening (N=2719; Table 3 below).

3.3 | Signal strengths in detected ≥CIN2

When considering all 4997 samples, HPV-positive ≥CIN2 with normal cytology had weaker median signal strengths than HPV-positive ≥CIN2 with abnormal cytology ($P<.001$ for HC2 and cobas, and $P<.002$ for APTIMA; Table 3 above). For HC2, the median RLU/CO was 25.6 (IQR: 8.6-67.2) in normal and 124.2 (IQR: 27.8-439.8, $P\leq.001$) in abnormal cytology. For cobas, the median Ct values were 30.4 (IQR: 27.9-32.9) and 26.8 (IQR: 23.9-30.9, $P\leq.001$), respectively. For APTIMA, the median S/CO values were 10.7 (IQR: 8.4-11.7) and 11.2 (IQR: 10.6-16.9, $P\leq.01$), respectively.

The differences in HPV signal strengths between normal and abnormal cytology associated with ≥CIN2 were also seen for samples from primary screening but were significant only for APTIMA ($P=.02$; Table 3 below). The primary screening population as a whole (Table 4), where the majority of ≥CIN2 in our population had abnormal cytology, also showed weaker signal strengths than the referral population with abnormal cytology, although the differences were not significant. Referral population samples with normal cytology (Table 5) were fairly similar to primary screening samples with normal cytology (Table 3 below).

3.4 | Concordance between HPV assays in detecting HPV infections and ≥CIN2

Among all 4997 samples, 1343 (29%) of 4630 samples with normal cytology had a positive test result on one or more HPV assays, but in 902 (67%) of the 1343 samples, the assays showed discordance (Table 6). For samples with abnormal cytology, the assays were discordant in 84/317 (26%). Hence, the relative risk of discordance in detecting HPV infections was significantly lower in abnormal compared to normal cytology: 0.39 (0.26/0.67, 95% CI: 0.33-0.48). In samples from primary screening (N=2846), the relative risk of assay discordance in detecting HPV infections was 0.42 (95% CI: 0.31-0.56) for abnormal vs normal cytology. The risk of assay discordance

TABLE 2 Distribution of signal strengths in HPV-positive women, by population and HPV assay

	All samples (N=4997)			Primary screening at 30-65 y (N=2846)		
	HC2	cobas	APTIMA	HC2	cobas	APTIMA
N testing positive (%)	1021 (20)	1343 (27)	837 (12)	335 (12)	464 (16)	270 (9)
Signal strengths in HPV-positive women						
Range	1.0-3100.5	11.8-40.5	0.5-63.9	1.0-2818.9	12.2-40.4	0.5-31.8
Mean (SD)	166.7 (380.1)	32.0 (5.4)	10.3 (6.2)	135.5 (329.5)	32.9 (5.4)	9.4 (5.5)
Median (IQR)	19.5 (4.3-132.6)	32.4 (27.9-36.7)	10.6 (7.1-11.8)	13.5 (3.7-97.0)	33.8 (28.7-37.9)	10.1 (6.0-11.3)

HC2, Hybrid Capture 2; HPV, Human Papillomavirus; IQR, interquartile range; SD, standard deviation.

TABLE 3 Median signal strengths (IQR) of detected HPV infections and HPV-positive ≥CIN2, by HPV assay and cytology test result at baseline

	HPV assay						
	HC2		cobas		CLART	APTIMA	
Baseline cytology result	Positive test result (%)	Md signal strength (IQR)	Positive test result (%)	Md signal strength (IQR)	Positive test result (%)	Positive test result (%)	Md signal strength (IQR)
All women (N=4997)							
HPV infections							
Normal (N=4630)	721 (16)	11.0 (3.3-52.8)	1061 (23)	33.5 (29.6-37.5)	992 (21)	581 (13)	10.2 (5.8-11.3)
Abnormal (N=367)	300 (82)	124.2 (22.8-506.9)	282 (77)	26.9 (23.7-31.3)	269 (73)	256 (70)	11.1 (9.4-15.5)
P[a]	–	<.001	–	<.001	–	–	<.001
≥CIN2							
Normal (N=57)	52 (91)	25.6 (8.6-67.2)	56 (98)	30.4 (27.9-32.9)	55 (96)	44 (77)	10.7 (8.4-11.7)
Abnormal (N=118)	114 (97)	124.2 (27.8-439.8)	113 (96)	26.8 (23.9-30.9)	112 (95)	106 (90)	11.2 (10.6-16.9)
P[a]	–	<.001	–	<.001	–	–	<.002
Primary screening at 30-65 y (N=2846)							
HPV infections							
Normal (N=2719)	249 (9)	9.4 (2.6-46.5)	385 (14)	34.5 (30.0-38.4)	377 (14)	199 (7)	9.2 (4.2-11.2)
Abnormal (N=127)	86 (68)	83.2 (10.0-330.7)	79 (62)	27.2 (25.2-33.0)	76 (60)	71 (56)	10.8 (9.6-12.9)
P[b]	–	<.001	–	<.001	–	–	<.001
≥CIN2							
Normal (N=12)	9 (75)	24.9 (7.8-88.8)	12 (100)	32.1 (29.1-34.0)	12 (100)	9 (75)	10.7 (8.3-11.8)
Abnormal (N=38)	37 (97)	92.6 (31.8-264.8)	37 (97)	27.1 (24.8-33.5)	37 (97)	37 (97)	10.8 (10.4-11.5)
P[b]	–	.08	–	.10	–	–	.02

CIN, cervical intraepithelial neoplasia; HC2, Hybrid Capture 2; HPV, Human Papillomavirus; IQR, interquartile range; Md, median.
[a]Adjusted for age (<30 vs ≥30 y), reason for taking the sample (primary samples vs follow-up samples), type of infection (single vs multiple).
[b]Adjusted for type of infection (single vs multiple).

TABLE 4 Median signal strengths (IQR) for samples with detected HPV infections and HPV-positive ≥CIN2, by HPV assay and type of population (primary screening vs referral population with abnormal cytology)

	Detection of HPV infections, Md signal strength (IQR)			Detection of ≥CIN2, Md signal strength (IQR)		
Population type	HC2	cobas	APTIMA	HC2	cobas	APTIMA
Primary screening at 30-65 y (N=2846)	13.5 (3.7-97.1)	33.8 (28.6-37.9)	10.1 (5.9-11.3)	75.7 (14.4-216.9)	28.1 (25.1-33.6)	10.8 (10.3-11.5)
Referral population with abnormal cytology (N=367)	124.2 (22.8-506.9)	26.9 (23.7-31.3)	11.1 (9.4-15.5)	124.2 (27.8-439.8)	26.8 (23.9-30.9)	11.2 (10.6-16.9)
P[a]	<.001	<.001	<.001	.54	.14	.17

CIN, cervical intraepithelial neoplasia; HC2, Hybrid Capture 2; HPV, Human Papillomavirus; IQR, interquartile range; Md, median.
Note: From the 367 women in the referral population with abnormal cytology (at any age), 317 (86%) were HPV-positive on at least one HPV assay, and 118 (32%) had HPV-positive ≥CIN2.
[a]Adjusted for type of infection (multiple vs single).

in referral population samples with normal cytology (N=516) was similar to that in cytologically normal screening samples (N=2719), relative discordance: 0.92 (95% CI: 0.84-1.02; not tabulated).

Discordance between the four assays was less frequent in samples associated with ≥CIN2, 26% in normal and 15% in abnormal cytology when all 4997 samples were included. In primary screening, discordance was 33% in normal cytology and 11% in abnormal cytology. Although assay discordance in detecting ≥CIN2 was two to three times as frequent in normal than in abnormal cytology, the difference did not reach statistical significance in our data.

4 | CONCLUSIONS

In our study, HPV infections had weaker median signal strengths in normal than in abnormal cytology. Being in the range of approximately

TABLE 5 Median signal strengths (IQR) for samples with detected HPV infections and HPV-positive ≥CIN2 in 516 women from a referral population with normal cytology, ie, undergoing follow-up testing for a recent abnormality and current samples showing normal cytology, by HPV assay

	Md signal strength (IQR)
Detected HPV infections	
HC2	8.5 (2.7-40.9)
Cobas	34.5 (30.3-37.8)
APTIMA	10.5 (6.5-11.2)
Detected ≥CIN2	
HC2	9.6 (4.6-68.0)
Cobas	33.0 (27.8-33.4)
APTIMA	10.3 (3.5-10.8)

CIN, cervical intraepithelial neoplasia; HC2, Hybrid Capture 2; HPV, Human Papillomavirus; IQR, interquartile range; Md, median.
Note: In the Horizon study, there were in total 516 women with normal cytology in samples taken for follow-up of a recent abnormality. Of these, 200 (39%) had a positive test result on at least one HPV assay, and 10 (2%) had HPV-positive ≥CIN2.

100 RLU/CO on HC2 and ca. 6 Ct on cobas, the absolute differences in the median signal strengths were not trivial. Abnormal cytology samples also showed a higher degree of assay concordance than normal cytology samples in detecting HPV infections.

HPV testing detects more ≥CIN2 than cytology.[30] At least part of ≥CIN2 in normal cytology is clinically important. In the two-round screening data from randomised trials, treatment of cytology-normal/HPV-positive ≥CIN2 reduced the population's risk of developing cervical cancer.[1] In our study, most ≥CIN2 were detected by all HPV assays. Nonetheless, the lower median signal strengths and assay discordance in women with normal cytology suggested that cytology-normal ≥CIN2 are not only difficult to detect through cytology but might also be more difficult to detect with HPV assays than is the case with cytology-abnormal ≥CIN2.

HPV assays utilise various molecular targets and methodologies, which result in different test dynamics. Differences in assay characteristics and calibration may, as discussed previously,[2] explain why the assays return discordant HPV test results. Higher amounts of the target viral input in a sample and the associated stronger signals may, however, make all assays more likely to return a positive test result. We observed this pattern in women with cytological and histological abnormalities, who, in turn, also showed a higher likelihood of concordance between HPV assays than women without abnormalities.

Based on these data, we can infer that referral population studies focusing predominantly on women with abnormal cytology underestimate the differences between HPV assays that would become apparent in primary screening. With a view on the practicality of undertaking the necessary assay validation studies, an international expert group has developed non-inferiority testing guidelines for primary screening. These guidelines require that new assays be compared to HC2.[31] They call for the use of a relatively small number of samples from population-based cohorts screened and managed according to HC2 and cytology test results. Even although this is not always followed in research practice, our data support the guidelines' requirement to include samples from true primary screening populations with a representative distribution of normal and abnormal cytology. However, if primary screening cohorts are not available for study, our data suggest that including a representative proportion of normal-cytology samples taken for follow-up of recent abnormalities might offer a reasonable shortcut. In our study, these samples had fairly similar signal strengths as normal-cytology primary screening samples and showed similar assay discordance. Nevertheless, the use of normal-cytology follow-up samples as a substitute for normal-cytology screening samples needs further validation.

The strength of our data was that samples taken for different reasons originated from the same population. All cytology was read by the same laboratory under routine conditions, which were not influenced by the study. Moreover, unlike in some other split-sample studies evaluating several HPV assays,[10] cytology-normal/HPV-positive women were invited for repeated testing. Approximately 60% of these women had a follow-up, regardless of the assay that detected the HPV infection.[27] This proportion was consistent with the data from randomised controlled trials comparing HPV testing to cytology in primary cervical screening.[32] Finally, management according to

TABLE 6 Discordance between HPV assays, by cytology test result

	All women (N=4997)					Women with primary screening samples, 30-65 y (N=2846)				
	Abnormal cytology (N=367)		Normal cytology (N=4630)			Abnormal cytology (N=127)		Normal cytology (N=2719)		
Endpoint	Positive on ≥1 assay	Positive on <4 assays	Positive on ≥1 assay	Positive on <4 assays	RD (95% CI) for abnormal vs normal cytology	Positive on ≥1 assay	Positive on <4 assays	Positive on ≥1 assay	Positive on <4 assays	RD (95% CI) for abnormal vs normal cytology
HPV infections	317	84 (26%)	1343	902 (67%)	0.39 (0.33-0.48)	93	30 (32%)	558	433 (78%)	0.42 (0.31-0.56)
≥CIN2	118	18 (15%)	57	15 (26%)	0.58 (0.32-1.06)	38	4 (11%)	12	4 (33%)	0.32 (0.09-1.07)

CI, confidence interval; CIN, cervical intraepithelial neoplasia; HPV, Human Papillomavirus; RD, relative discordance.
Note: In women with normal cytology undergoing follow-up for a recent abnormality (N=516), 200 (39%) tested positive on at least one HPV assay. Of the 200 women, 143 (72%) did not test positive on all four assays. The concordance between the four assays in detecting HPV infections was similar as in women with normal cytology undergoing primary screening, relative discordance for normal cytology in follow-up vs primary screening: 0.92 (95% CI: 0.84-1.02).

HPV status was not limited to the results of a single assay, as women qualified for repeated testing if one or more HPV assays were positive. This allowed for a more comprehensive evaluation of the differences in detecting HPV-related abnormalities between the four assays than has been the case in several previous studies.

Nevertheless, this is a post-hoc analysis, and the study was not powered to detect differences between cytology-normal and cytology-abnormal CIN. In primary screening, only 12 out of 50 ≥CIN2 were associated with normal cytology. With larger numbers of screened women, the observed differences in the ability of the assays to detect ≥CIN2 could, plausibly, reach statistical significance. Studies from other populations are warranted to corroborate our findings.

In conclusion, differences between HPV assays tend to be underestimated when studied only in women with abnormal cytology. Therefore, experiences and conclusions drawn from studies without a representative sample of women with normal cytology should be only cautiously used as evidence for the use of new HPV assays in primary cervical screening.

ACKNOWLEDGMENTS

The authors would like to thank women for the participation in the study follow-up. The authors would, furthermore, like to acknowledge the Danish Genomica, Hologic, QIAGEN, and Roche representatives for enabling the collaboration between their parent companies, Copenhagen University Hospital Hvidovre and the University of Copenhagen.

CONFLICT OF INTEREST

M.R. and her former employer received honoraria from QIAGEN for lectures on her behalf. S.N. declares no conflict of interest. E.L. participated in meetings with Roche and Astra-Zeneca with fees paid to the University of Copenhagen and was an unpaid advisor to GenProbe and NorChip. S.P. received honoraria from Hologic for lectures. D.E. attended sponsored meetings and symposia, received honoraria from BD and Qiagen for lectures and is project manager on studies funded by BD Diagnostics. C.R. participated in a meeting with Roche with the fee paid to the University of Copenhagen. J.B. used to serve as a paid advisor to Roche and Genomica, and has received honoraria from Hologic/Gen-Probe, Roche, Qiagen, Genomica, and BD diagnostics for lectures. He is the principal investigator on studies funded by BD diagnostics. M.R., J.B., S.P., D.E., C.R., E.L. attended meetings with various HPV technology manufacturers. Hvidovre Hospital holds a recompense agreement with Genomica on a KRAS/BRAF diagnostic system.

REFERENCES

1. Ronco G, Dillner J, Elfström KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet*. 2014;383:524-532.
2. Rebolj M, Preisler S, Ejegod DM, Rygaard C, Lynge E, Bonde J. Disagreement between human papillomavirus assays: an unexpected challenge for the choice of an assay in primary cervical screening. *PLoS One*. 2014;9:e86835.
3. Iftner T, Becker S, Neis KJ, et al. Head-to-head comparison of the RNA-based aptima human papillomavirus (HPV) assay and the DNA-based hybrid capture 2 HPV test in a routine screening population of women aged 30 to 60 years in Germany. *J Clin Microbiol*. 2015;53:2509-2516.
4. Heideman DA, Hesselink AT, Berkhof J, et al. Clinical validation of the cobas 4800 HPV test for cervical screening purposes. *J Clin Microbiol*. 2011;49:3983-3985.
5. Boers A, Wang R, Slagter-Menkema L, et al. Clinical validation of the cervista HPV HR test according to the international guidelines for human papillomavirus test requirements for cervical cancer screening. *J Clin Microbiol*. 2014;52:4391-4393.
6. Lloveras B, Gomez S, Alameda F, et al. HPV testing by cobas HPV test in a population from Catalonia. *PLoS One*. 2013;8:e58153.
7. Wu R, Belinson SE, Du H, et al. Human papillomavirus messenger RNA assay for cervical cancer screening: the Shenzhen Cervical Cancer Screening Trial I. *Int J Gynecol Cancer*. 2010;20:1411-1414.
8. Ratnam S, Coutlee F, Fontaine D, et al. Aptima HPV E6/E7 mRNA test is as sensitive as hybrid capture 2 assay but more specific at detecting cervical precancer and cancer. *J Clin Microbiol*. 2011;49:557-564.
9. Szarewski A, Ambroisine L, Cadman L, et al. Comparison of predictors for high-grade cervical intraepithelial neoplasia in women with abnormal smears. *Cancer Epidemiol Biomarkers Prev*. 2008;17:3033-3042.
10. Cuzick J, Cadman L, Mesher D, et al. Comparing the performance of six human papillomavirus tests in a screening population. *Br J Cancer*. 2013;108:908-913.
11. Clad A, Reuschenbach M, Weinschenk J, Grote R, Rahmsdorf J, Freudenberg N. Performance of the APTIMA high-risk HPV mRNA assay in a referral population in comparison with Hybrid Capture 2 and cytology. *J Clin Microbiol*. 2011;49:1071-1076.
12. Mesher D, Szarewski A, Cadman L, et al. Comparison of human papillomavirus testing strategies for triage of women referred with low-grade cytological abnormalities. *Eur J Cancer*. 2013;49:2179-2186.
13. Lapierre SG, Sauthier P, Mayrand MH, et al. Human papillomavirus (HPV) DNA triage of women with atypical squamous cells of undetermined significance with cobas 4800 HPV and Hybrid Capture 2 tests for detection of high-grade lesions of the uterine cervix. *J Clin Microbiol*. 2012;50:1240-1244.
14. Fornari D, Rebolj M, Bjerregård B, et al. Hybrid Capture 2 and cobas human papillomavirus assays perform similarly on SurePath samples from women with abnormalities. *Cytopathology*. 2016;27:249-260.
15. Innamaa A, Dudding N, Ellis K, et al. High-risk HPV platforms and test of cure: should specific HPV platforms more suited to screening in a 'test of cure' scenario be recommended? *Cytopathology*. 2015;26:381-387.
16. Pista A, Verdasca N, Oliveira A. Clinical performance of the CLART human papillomavirus 2 assay compared with the hybrid capture 2 test. *J Med Virol*. 2011;83:272-276.
17. Reuschenbach M, Clad A, von Knebel Doeberitz C, et al. Performance of p16INK4a-cytology, HPV mRNA, and HPV DNA testing to identify high grade cervical dysplasia in women with abnormal screening results. *Gynecol Oncol*. 2010;119:98-105.
18. Cuschieri K, Cubie H, Graham C, et al. Clinical performance of RNA and DNA based HPV testing in a colposcopy setting: influence of assay target, cut off and age. *J Clin Virol*. 2014;59:104-108.
19. Stoler MH, Wright TC Jr, Cuzick J, et al. APTIMA HPV assay performance in women with atypical squamous cells of undetermined significance cytology results. *Am J Obstet Gynecol*. 2013;208:144.e1–8.

20. Stoler MH, Wright TC Jr, Sharma A, Apple R, Gutekunst K, Wright TL. Highrisk human papillomavirus testing in women with ASC-US cytology: results from the ATHENA HPV study. *Am J Clin Pathol*. 2011;135:468-475.
21. Rebolj M, Bonde J, Preisler S, Ejegod D, Rygaard C, Lynge E. Differential detection of human papillomavirus genotypes and cervical intraepithelial neoplasia by four commercial assays. *J Clin Microbiol*. 2016;54:2669-2675.
22. Preisler S, Rebolj M, Untermann A, et al. Prevalence of human papillomavirus in 5,072 consecutive cervical SurePath samples evaluated with the Roche cobas HPV real-time PCR assay. *PLoS One*. 2013;8:e59765.
23. Bonde J, Rebolj M, Ejegod DM, Preisler S, Lynge E, Rygaard C. HPV prevalence and genotype distribution in a population-based split-sample study of well-screened women using CLART HPV2 human papillomavirus genotype microarray system. *BMC Infect Dis*. 2014;14:413.
24. Rebolj M, Preisler S, Ejegod DM, Bonde J, Rygaard C, Lynge E. Prevalence of human papillomavirus infection in unselected SurePath samples using the APTIMA HPV mRNA assay. *J Mol Diagn*. 2013;15:670-677.
25. Rebolj M, Lynge E, Ejegod D, Preisler S, Rygaard C, Bonde J. Comparison of three human papillomavirus DNA assays and one mRNA assay in women with abnormal cytology. *Gynecol Oncol*. 2014;135:474-480.
26. Rebolj M, Bonde J, Ejegod D, Preisler S, Rygaard C, Lynge E. A daunting challenge: human papillomavirus assays and cytology in primary cervical screening of women below age 30 years. *Eur J Cancer*. 2015;51:1456-1466.
27. Rebolj M, Bonde J, Preisler S, Ejegod D, Rygaard C, Lynge E. Human papillomavirus assays and cytology in primary cervical screening of women aged 30 years and above. *PLoS One*. 2016;11:e0147326.
28. Preisler S, Rebolj M, Ejegod DM, Lynge E, Rygaard C, Bonde J. Cross-reactivity profiles of hybrid capture II, cobas, and APTIMA human papillomavirus assays: split-sample study. *BMC Cancer*. 2016;16:510.
29. Bjerregaard B, Larsen OB. The Danish pathology register. *Scand J Public Health*. 2011;39:72-74.
30. Rijkaart DC, Berkhof J, Rozendaal L, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol*. 2012;13:78-88.
31. Meijer CJ, Berkhof J, Castle PE, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *Int J Cancer*. 2009;124:516-520.
32. Rebolj M, Lynge E. Incomplete follow-up of positive HPV tests: overview of randomised controlled trials on primary cervical screening. *Br J Cancer*. 2010;103:310-314.
33. Andersen JS, de Fine Olivarius N, Krasnik A. The Danish National Health Service register. *Scand J Public Health*. 2011;39:34-37.
34. Denmark Statistics. Internal migration between municipalities by sex, age and municipality. http://www.statistikbanken.dk/statbank5a/default.asp?w=1366. Accessed April 29, 2016.

How to cite this article: Rebolj M, Njor S, Lynge E, et al. Referral population studies underestimate differences between human papillomavirus assays in primary cervical screening. *Cytopathology*. 2017;28:419-428. https://doi.org/10.1111/cyt.12451

APPENDIX

FOLLOW-UP RECOMMENDATIONS FOR WOMEN WITH CYTOLOGICAL ABNORMALITIES

As the study was nested into routine screening, women with cytological abnormalities were managed according to the routine recommendations valid at the time of the study. The study did not affect these processes. An immediate referral for colposcopy was recommended to women with high-grade squamous intraepithelial lesions (HSIL) or worse, atypical glandular cells, atypical squamous cells—cannot exclude HSIL and HC2-positive ASCUS at age ≥30 years. Other women with ASCUS and women with low-grade squamous intraepithelial lesions (LSIL) were recommended for repeated cytology-based testing. At follow-up testing, they were referred for colposcopy in case of repeated abnormalities. Among women undergoing primary screening at age 30-65 years, 92% had follow-up in 2.5 years after the baseline testing.[25]

FOLLOW-UP RECOMMENDATIONS FOR CYTOLOGY-NORMAL/HPV-POSITIVE WOMEN

At the time of the study, HPV-based primary screening was not recommended in Denmark, and cytology-normal/HPV-positive women had to be invited for follow-up for research purposes. Women were invited for repeated cytology and HPV testing in approximately 18 months after the baseline testing, independent of which HPV assay(s) detected their infection. At follow-up, women were recommended for colposcopy if they had ≥ASCUS or a positive HC2 test result. The ethical committee of the Capital Region approved of this follow-up and women provided written informed consent. As the study was nested into routine practice, women may have also had follow-up other than study testing. Combined, approximately 60% of the women had follow-up in 2.5 years after the baseline testing.[27]

DEFINITION OF PRIMARY SCREENING SAMPLES

In the Patobank,[29] the reason for taking the sample is not registered systematically. We used an algorithm to determine which samples were most likely taken for primary screening, ie, not for follow-up of a recent abnormality. Primary samples were defined as those without a:

- previous histological diagnosis of cervical cancer,
- histologically-confirmed cervical intraepithelial neoplasia (CIN) of any grade in ≤3 years,
- ASCUS cytology or non-CIN cervical histology in ≤15 months,
- more severe cytological abnormality, inadequate cytology or a positive HPV test result in ≤12 months.

Reflecting routine practice, primary samples included a small proportion of samples taken for investigation of symptoms. The Patobank data were retrieved from 1 January 2000 onwards.

Registration of cervical cytology and histology in the Patobank became nationally complete only gradually. Most screening samples are taken by privately practicing general practitioners, whereas private gynaecology practices are available particularly in the more urbanised areas. Cytology and histology samples can be read by pathologists working in hospital laboratories (where the registration in the Patobank has been complete since 1997)[29] and by privately practicing pathologists. Registration of samples in the national Patobank is now mandatory also for private pathologists, although the year in which this was implemented depended on the geographical area. Since 1990, cervical screening and histology samples taken by, or evaluated by privately practicing specialists, have also been registered in the National Health Service Register.[33] As this is a reimbursement register, it is considered highly complete but does not contain diagnostic information. For each woman, and for cytology separately from histology, we compared the numbers of registered cervical samples in the Patobank and the National Health Service Register until the end of 2010. All duplicates were excluded. We defined the start of (reasonably) complete registration in the Patobank as the year in which it started to include at least 85% of all samples. For cervical cytology, this analysis suggested that for the period 2000-2010, approximately 7% of all samples were not registered in the Patobank; this was 11% for cervical histology. Nevertheless, Copenhagen County had highly incomplete registration of cytology before 2005, and Sønderjylland County before 2001. Our threshold for considering the registration reasonably complete was reached for histology in Frederiksberg Municipality in 2001, Sønderjylland and Frederiksborg Counties in 2003, Storstrøm County in 2006, and Copenhagen County in 2009. The remaining counties reached the threshold before 2000.

Women from (1) Copenhagen Municipality and (2) Frederiksberg Municipality, the catchment areas of our laboratory at the time of the study, have had highly complete registration of cytology since (1) 1991 and (2) 1993, and highly complete registration of histology since (1) 1998 and (2) 2001. Incomplete registration elsewhere in Denmark will affect our analyses; however, the impact was most likely limited as migration to the two municipalities from elsewhere in Denmark was infrequent for women aged 30-65 years (estimated at <3% for the year 2005).[34]

HPV TESTING PROTOCOLS IN THE HORIZON STUDY

Processing of samples and assay instrumentation

The study protocol, sample storage, and assay testing protocols were agreed upon with all manufacturers before the study. All instrumentation and software were used as supplied and maintained by the manufacturers.

Baseline

These testing protocols were published, with minor textual revisions, previously.[2]

HC2

We used cytology post-quot material that remained from the cytology procedure. As part of the cytology processing, post-quot material was diluted approximately 1:1 in SurePath. DNA was either denatured prior to testing by pre-treating manually in line with the manufacturer's CE-IVD protocol, or was isolated and purified using the DSP AXpH DNA kit on QIASymphony SP (Qiagen, Hilden, Germany). Testing was undertaken on automated Rapid Capture System (RCS; Qiagen, Gaithersburg, MD, USA). A minority of samples used for routine HC2 triage of women with ASCUS at age ≥30 years were denatured and tested manually.

Cobas

1 mL of the diluted material was aliquoted into a 13 mL round bottom test tube (Sarstedt, cat. no NC9018280), stored at 2-8°C until testing. No pre-treatment of SurePath samples was required. Extraction of DNA was undertaken on cobas ×480, and amplification and detection of high-risk HPV DNA on cobas z480 analyser. Fluorescent TaqMan probes were used for detection of the amplicons during polymerase chain reaction (PCR) cycles. Amplification and detection of the 330-bp β-globin was used as an internal control of the testing processes.

CLART

1 mL of the diluted SurePath sample was spun down for 5 minutes at 14 000 revolutions per minute, with supernatant removed and cell pellet re-suspended in a mix of 180 μL phosphate buffered saline (10× conc. pH 7.4, Pharmacy product) and 20 μL Proteinase K (recombinant, PCR Grade, Roche Diagnostics, Rotkreuz, Switzerland). Samples were then vortexed and incubated for 1 hour at 56°C and 1 hour at 90°C. HPV DNA was purified using MagNa Pure LC 96 and MagNA Pure LC 32 instruments (Roche Diagnostics) with MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche Diagnostics). PCR amplification was performed using the CLART HPV2 Amplification kit (Genomica). 5 μL of purified DNA were used for the PCR amplification. Prior to visualisation, the PCR products were denatured at 95°C for 10 minutes. Visualisation was performed using 10 μL of the denatured PCR products on the CLART microarray. Hybridisation between the amplicons and their specific probes on the microarray resulted in formation of an insoluble precipitate of peroxidase when adding a Streptavidin conjugate that binds to the biotin-labeled PCR products. The precipitate was analyzed automatically on the Clinical Array Reader (Genomica).

APTIMA

1 mL of the diluted sample was aliquoted into an APTIMA Specimen Transfer Tube containing 2.9 mL of buffered solution (Hologic/GenProbe). Samples were treated with proteinase K prior to testing,

using the Pace 2 Fast Expression Kit containing 1 mL diluent and lyophilised reagent (all from Hologic/Gen-Probe). 100 μL of the reconstituted proteinase K was added to each Specimen Transfer Tube and incubated at 65°C for 2 hours. The treated specimen tube was stored at 2-8°C until testing. Testing was performed on the PANTHER platform.

Follow-up testing for women with cytology-normal/HPV-positive test results at baseline

Testing on HC2 for all follow-up samples was performed on the cytology post-quot material with manual DNA denaturation followed by testing on the RCS.